View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by edoc

Development of a simple Microarray for Genotyping HIV-1 Drug Resistance Mutations in the Reverse Transcriptase Gene in rural Tanzania

Pax Masimba[1,2,4*], Janet Gare[1,2*&], Thomas Klimkait[2,3], Marcel Tanner[1,2], Ingrid Felger[1,2]

*Shared first authorship

1 Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

2 University of Basel, CH-4003 Basel, Switzerland

3 Department Biomedicine, Haus Petersplatz, University of Basel, Switzerland

4 Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania

& Current affiliation: Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea

ABSTRACT

Objective The success of antiretroviral therapy (ART) of HIV-1 is compromised by development of drug resistance (DR) due to mutations in viral target genes. Monitoring of these DR mutations will help to avoid continuation of ineffective therapies and contribute to optimization of ART. In Tanzania, molecular analysis of DR is currently limited owing to high cost. Therefore, a simple, inexpensive and robust tool for DR genotyping was developed based on microarray technology and was restricted to 25 DR mutations most relevant for the locally available ART regimen.

Methods The reverse transcriptase gene fragment was reverse transcribed and amplified by Reverse Transcription-Polymerase Chain Reaction (PCR). Primers for mini-sequencing were designed based on alignments of most prevalent local HIV-1 variants. Tagged primers were extended by fluorochrome-labeled dideoxynuclotide triphosphate (ddNTPs) to indicate the Single Nucleotide Polymorphism (SNP) allele of the sample tested, followed by hybridization on treated microarray slides. Images were analyzed with a laser scanner and genotype calling was performed using in-house developed software. The microarray was validated with four cloned HIV-1 genome fragments from Swiss HIV-1 Cohort and 102 HIV-1 sequences amplified from Tanzanian target population (field samples) and the results were concordant with the Sanger sequencing SNP profile in 92.7% of 2550 SNP data points compared. Lack of signals in small number of SNPs was due to either failure in extension reaction or

hybridization owing to excessive mismatches between PCR product and extension primer.

Conclusion Our study demonstrates the feasibility of hybridization-based genotyping of drug resistance mutations of HIV, even though the microarray, which was designed for population studies, could achieve an assignment of only 92% for the individual SNPs in the tested samples.

INTRODUCTION

Development of HIV-1 drug resistance has been a major obstacle in the long-term success of Antiretroviral Therapy (ART) for HIV-1 patients (Shafer *et al*., 2000). One of the main factors accounting for the development of drug resistance is the emergence of mutations in the reverse transcriptase gene, which is the major HIV-1 drug target (Bean, 2005).

A number of phenotypic and genotypic assays have been used to detect HIV-1 drug resistance (DR) mutations (Grant and Zolopa, 2009). Phenotypic assays measure directly the extent to which an antiretroviral drug inhibits HIV-1 replication in-vitro, and determine an increase in the inhibitory concentration that is required to inhibit in-vitro replication by 50 percent compared with virus replication in the absence of drug. Results are reported as fold-change in drug susceptibility of the patient-derived virus sample compared with a laboratory reference strain of HIV-1 (MacArthur, 2009). Phenotypic testing reflects the net effect of HIV-1 mutations on susceptibility to each tested drug and has advantages in patients with complex mutation patterns (Hirsch *et al*., 2008). Genotypic HIV resistance assays assess known mutations associated with drug resistance. These assays involve the detection of genomic HIV-1 mutations in regions that are targeted by the current ARV drugs, mainly protease, reverse transcriptase, integrase, and envelope glycoprotein 41. Following amplification of the genes of interest by reverse transcription-PCR, the amplicons can further be processed by two different approaches: either direct sequencing or hybridization-based methods. While sequencing determines the full nucleotide sequence of the selected gene for analysis,

hybridization targets only specific SNPs of interest. Compared to phenotypic assays, genotypic testing has the advantages of shorter turn-around time and lower cost (Hirsch *et al*., 2008).

While in developed countries both genotypic and phenotypic assays are routinely used for monitoring HIV-1 patients at initiation of ART and in case of suspected drug resistance, these assays are currently rarely available in developing countries due to their prohibitive costs. In view of the rapidly increasing need for molecular monitoring of the prevalence and spread of DR also in resource-poor settings, an additional molecular tool for robust and affordable DR genotyping is warranted. While nucleotide sequencing certainly remains the gold standard for molecular detection of DR-SNPs, surveillance of population samples and molecular epidemiological research project in developing countries like Tanzania could greatly benefit from a simple and robust tool to determine the limited number of DR-SNPs that are most critical for a specific location and available drugs and drug combinations. We therefore investigated the option to genotype multiple SNPs using a standard microarray platform.

A related microarray platform had been developed previously in our laboratory for detection of SNPs in drug resistance marker genes of malaria parasites (Crameri *et al*., 2007). For this application the mini-sequencing principle was validated with highly specific base calling and parallel genotyping of many SNPs at a time (Syvänen, 1999). We intended to adopt this platform for genotyping local HIV-1 variant in Ifakara, a site in rural Tanzania. For this proof-of-concept the HIV-1 microarray was restricted to

mutations associated with resistance to the available reverse transcriptase inhibitors (RTI) in use in the Ifakara HIV-1 cohort, i.e. stavudine, lamivudine, zidovudine, abacavir, didanosine, nevirapine, and efavirenz.

Due to the intrinsic high error rate of the HIV-1 polymerase extensive polymorphisms are generally observed among circulating HIV-1 strains. Therefore we focused the design of primers on local subtypes and strains occurring in the study area. Our intention was to provide an affordable option for monitoring HIV-1 drug resistance in Ifakara, Tanzania, by genotyping via microarray; this would permit a highly multiplexed SNP analysis in a single run, requiring little hand-on-time and resources as demonstrated in the malaria chip-project (Crameri *et al*., 2007).

MATERIALS AND METHODS

Patients and plasma samples

Samples used in this study were obtained from HIV-1 patients from the KIULARCO HIV cohort of Ifakara, Tanzania. KIULARCO was established in 2004 in Kilombero and Ulanga districts, Morogoro region, Southern Tanzania, for the purpose of implementing care and treatment of HIV/AIDS patients according to Tanzania National AIDS Control Care Programme (NACP) and to conduct – alongside with offering these essential services - applied research on minimal essential care and monitoring for HIV patients in rural resource-poor, peripheral settings. Patients were enrolled at the Chronic Disease Center Ifakara (CDCI) affiliated with Ifakara Health Institute and St. Francis Referral

Hospital (SFRH), the main district hospital providing treatment and care for a population of more than 600,000 and for 30,000 people living with HIV/AIDS (Mossdorf *et al*., 2011).

Plasmid-cloned HIV-1 genes used for validation of the microarray stemmed from anonymous clinical HIV-1 samples collected in Switzerland.

Ethical Considerations

This study was approved by ethics review bodies of Tanzania, the Ifakara Health Institute (IHI) Institutional Review Board, and the Medical Research Coordination Committee of the National Institute for Medical Research (NIMR) through the Tanzania Commission for Science and Technology (COSTECH) and the Ethics Committee of Canton Basel (EKBB), Switzerland. Patients taking part in this study had given their informed written consent.

RNA extraction, Reverse Transcription and PCR

Viral RNA was extracted from plasma with either the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) or Macherey-Nagel NucleoSpin RNA Virus Kit (Macherey-Nagel GmbH & Co KG, Neumann-Neander, Germany) using the manufacturers' protocols.

Reverse transcription was performed using specific primer RT2, AffinityScript RT Buffer (500 mM TrisHCl pH 8.3, 750 mM KCl, 30 mM MgCl2), 2 µL of 100 mM DTT (Stratagene, North Torrey Pines Road La Jolla, CA), 0.8 µL dNTP mix (25 mM each dNTP), 1 µL of a RNase Inhibitor, RNase Out (40 U/µL), 1 µL AffinityScript Multiple Temperature Reverse Transcriptase, 1 µL specific Primer RT2 (5'-GATAAGCTTGGGCCTTATCTATTCCAT-3`), (10 µM), HPLC purified, and 9.5 µL RNA

solution. Reverse transcription was performed with the following thermal conditions: 42°C for 35 min, 55°C for 25 min, 70°C for 15 min and 5°C for 15 min. All primers used in this work were synthesized by Eurofins MWG Operon, Ebersberg, Germany.

Primary PCR (pPCR) was done using Advantage cDNA Polymerase according to the supplier's protocol (Clontech Laboratories, Inc.Mountain, View, USA) with some modifications. Reverse and forward primers RT2 and D1818 (5'-AGAAGAAATGATGACAGCATGTCAGGGAGT-3`) were used. The pPCR mix contained 5 µL 10x Advantage buffer (Clontech), 10 µL dNTP mix (2 mM), 2 µL reverse primer RT2 (10 µM), 2 µL forward primer D1818 (10 µM), 1 µL Advantage Polymerase (5U/µL) and 4 µL of cDNA. Reaction profile was 94˚C for 2 min, 94˚C for 20 sec, 47˚C for 20 sec, 68˚C for 2 min, 30 cycles followed by a final elongation step at 68˚C for 5 min.

A fragment of 645 bp spanning positions 23 – 236 in the reverse transcriptase gene was then amplified by nested PCR (nPCR). The nPCR mix contained 5 µL 10x Pfu buffer (Promega Corporation, Woods Hollow Road, Madison, WI USA), 10 µL dNTP mix (2 mM), 2 µL forward primer JG103 5`-AACAATGGCCATTGACAGAA[I-Q]-3` (10 µM), 2 µL reverse primer JG202 5`-TCAGGATGGAGTTCATAICCCA-3` (10 µM), 0.7 µL FIREPol Polymerase (3U/µL), 0.1 µL Pfu Polymerase (3U/µL) and 2 µL pPCR product. Thermocyling conditions were: 94˚C for 2 min, followed by 30 cycles of 94˚C for 15 sec, 47˚C for 15 sec, 72 ˚C for 2 min and a final elongation step at 72˚C for 5 min.

Nucleotide Sequencing

Direct Sanger sequencing of PCR products from field samples and cloned HIV-1 genome fragments was performed as previously described (Masimba et al., 2013). In brief, sequencing was performed either in our lab or by the commercial supplier Macrogen, South Korea. The in-house protocol used forward primers JG103 (5'-AACAATggCCATTgACAgAA[I-Q]-3') or PMF (5'-AACTCAAGACTTT TGGGAAGT-3') or

171 the reverse primers JG202 (5'-TCAggATggAgTTCATAICCCA-3') or PMR (5'-
172 TTGTCATGCTACTCTGGAATA-3'). PMF and PMR are centrally located sequencing
173 primers for the reverse transcriptase gene. Sequences were aligned using SeqScape
174 Software Programme Version 2.7 (AB, Applied Biosystems, Foster City, CA). The
175 sequence accession numbers are KC537065–KC537290.
176
177 **Design of Extension Primers, Tags and Anti-Tags**
178 Extension primers were designed for 25 prioritized DR-SNPs in the reverse
179 transcriptase gene. Per SNP one or more extension primers were designed using a
180 Clustalw1 alignment of 126 sequences from HIV-1 samples of patients from the
181 KIULARCO cohort in Ifakara, Tanzania. Extension primers were designed in either the
182 forward or reverse direction to maximize sequence conservation between the designed
183 primer and the variety of template sequences. Single base extension (SBE) software (6)
184 was used to design a set of 100 tags and anti-tags with the following parameters: length
185 17-25 bp, melting temperature (Tm) 53-62°C, homodimer temperature 40°C. One
186 individual tag was added to the 5` end of a single extension primer, whereby the SBE
187 program was used to select the optimal tag/extension primer pairs by assessing Tm and
188 potential for hairpin formation, homodimer and heterodimer formation. For spotting on
189 microarrays 55 anti-tags, i.e. reverse complement of the tags selected for the extension
190 primers, were synthesized via a C7 aminolinker for covalent coupling to the aldehyde
191 glass slide. Oligonucleotides used as extension primers, anti-tags, as well as Cyanine 5
192 (Cy5)- and Cyanine 3 (Cy3)-prelabeled oligonucleotides used as spotting controls and

one additional Cy5-labeled hybrization control were all purchased from Eurofins. Labeled ddNTPs were purchased from Perkin Elmer, Schwerzenbach, Switzerland.

Array design and production

55 anti-tags plus 2 Cy3- and Cy5-prelabeled spotting controls were spotted at the Center of Integrative Genomics, University of Lausanne, Switzerland, on Arrayit aldehyde-coated slides with a 12 well mask (Supermask 12 Super Aldehyde Slides purchased from Anopoli Biomedical Systems, Eichgraben, Austria). Oligonucleotides were dissolved as a 10x stock (500µM) in 180 mM phosphate buffer pH 8.0 and spotted after dilution in spotting buffer (3xSSC buffer with 1.5 M betaine) at a concentration of 50 µM (labeled spotting controls at a concentration of 0.5 µM). The array was printed in triplicate per masked well. After printing, slides were kept on a chamber at 50% relative humidity and baked the next day at 80°C for 90 min.

Prior to hybridization slides were pre-incubated at 80°C for 90 min, followed by two washing steps for 2 min in 0.2% SDS and three times for 2 min in distilled water and was dried by centrifugation at 800 rpm for 5 min. Slides were reduced in 50mM triethanolamine titrated with boric acid to pH 8.0 at 50°C for 30 min, washed three times with 0.2% SDS for 1 min, then two times in distilled water for 1 min, and finally dried by centrifugation for 5 min at 800 rpm. Slides were then kept at room temperature in a dry, clean and dark place until used in hybridization experiments, usually on the same day or within one week.

Primer Extension and hybridization

Prior to the primer extension reaction, nested PCR products were subjected to a Shrimp Alkaline Phosphatase (SAP) digest (Amersham Biosciences, Freiburg, Germany) to eliminate all non-incorporated nucleotides. This reaction was carried out as previously described (Shafer *et al.*, 2009). Primer extension with Cy3- and Cy5-labeled dideoxynucleotide triphosphates (Cy3-ddNTPs and Cy5-ddNTPs from Perkin Elmer) was carried out as described previously (Crameri *et al*., 2007). As the scanner supported only dual fluorescence measures, two extension reactions were performed with different permutations of Cy3- and Cy5- labeled ddNTPs. Table 1 shows the composition of both reaction mixes and indicates the required reaction mix for each extension primer. These two combinations of differentially labeled ddNTP were sufficient to differentiate all wild type from mutant alleles. Extension products from both reactions were combined before denaturation and hybridization performed as described previously (Crameri *et al*., 2007) with the modification that hybridization was performed at 55^0C for two hours.

Washing

After hybridization, slides were washed at room temperature (20^0C) in 3 consecutive buffers (temperature of the washing buffer was set at around 25^0C). One wash round consisted of 2X SSC + 0.2% SDS for 3 min., followed by 2X SSC for 2 min and finally 2X SSC + 2% Ethanol for 1 min. The number of rounds depended on background fluorescence and fluorescence intensity of the spots. To adjust washing conditions to the background intensity, a slide was quickly dried with compressed air after each round

and then pre-scanned. Further washing rounds were added if needed until the background fluorescence was satisfactorily removed without compromising signal intensity. Usually two to three rounds were sufficient.

Image and Genotype Scoring

After drying slides were scanned in a GenePix® microarray scanner 4100A (Axon Instruments, Genepix, USA) and images were stored as Tagged Image File Format (TIFF) file. Images were interpreted by running the GenePix software in combination with a file containing the array layout and a custom-made script. All spots with pre-labeled tags or anti-tags gave strong signals at their defined locations and could thus be used to position the array. The data retrieved was stored in a GenePix Result (GPR) file, which was transferred to an in house generated receiver operating characteristic (ROC) Classifier program for SNP calling. This program evaluated and scored the triplicate hybridization signals for each SNP into wild type or mutant based on threshold values from a set of positive (triplicate spotting and hybridization controls and negative (unused anti-tags) controls present on each slide.

Cloned HIV plasmids

Four Cloned fragments of the HIV-1 genome were used for microarray validation. These fragments, derived from anonymous Swiss HIV-1 Cohort samples and cloned in puc18 plasmid were made available from DBM, Haus Petersplatz, University of Basel (Fehr *et al*., 2011). After transformation the individual bacterial colonies were picked and plasmid DNA was extracted using QIAprep Miniprep Spin columns (Qiagen, Germany)

according to the manufacturer's instructions. Each cloned fragment represented a single HIV-1 reverse transcriptase gene variant suitable for assessing background hybridization and test validation. Plasmid inserts were of HIV-1 subtypes A, AE/A or C. The subtypes of cloned fragments were chosen to represent African HIV-1 subtypes.

RESULTS

Design of extension primers, tags and anti-tags

For 25 SNPs in the reverse transcriptase gene, a total of 51 extension primers (1 to 8 extension primers per SNP) were designed to compensate genetic diversity in the targeted sequence. Extension primers, tags and anti-tags are listed in **Supplementary Table 1**. Prior to the hybridization, all extension primers were tested by individual PCRs involving one of the extension primers plus either the forward or reverse primers normally used in our nPCR. As templates, four cloned reverse transcriptase gene fragments (subtypes A, CRF_02/AE and twice C) were used as well as 102 reverse transcription-PCR products from Tanzanian HIV-1 patients. All extension primers yielded DNA fragments of the expected size, indicating a sufficient degree of sequence conservation between these primers and the different templates tested (**Supplementary Figure 1**).

Array design and spotting

SBE software was used to select 55 oligonuclotide anti-tags and 2 spotting controls, the latter being produced with Cy3- or Cy5- fluorescent label at their 3' end. One anti-tag

was reserved for a hybridization control, for which the Cy5- pre-labeled tag was added to the extended primers prior to hybridization. Unused tags and printed buffer spots were used as negative controls. The array printed in Arrayit slides consisted of the 55 anti-tags and controls in a 14 x 14 spot lay out (**Figure 1**), thus generating 3 data points for each position per sample. The separation of each slide by a hydrophobic mask into 12 separate reaction areas or wells permitted parallel analysis of 12 patient samples per array. In contrast to the Cy3-control the Cy5-labeled spotting control was found to decay rapidly during storage of slides. We therefore used a hybridization control that also carried a Cy5- label for verifying hybridization success as well as for correct positioning of each array with the help of fixed and strong signals of control spots. The results obtained from all pre-labeled controls confirmed a good reproducibility of the hybridization on microarray.

Optimization of washing

Despite systemic tag/anti-tag design, establishing of optimal wash conditions for hybridized slides was initially a challenge in generating optimal signal intensities for all spots of an array. The number of wash steps depended on the background fluorescence and the spot intensity on each particular slide and was therefore adapted after a pre-scan after each washing round. The majority of slides w washed three times, with each round consisting of 2X SSC + 0.2% SDS for 3 min, followed by 2X SSC for 2 min and finally 2X SSC + 2% Ethanol for 1 min. Minor differences in slide pretreatment conditions and duration of storage could have contributed to these differences.

Validation of microarray

For validating the microarray's accuracy in base calling we compared the SNP profile of a sample generated by microarray with the same SNPs derived from Sanger sequencing of the respective cloned HIV-1 genome fragment. In contrast to PCR fragments from field samples, which may represent mixed populations of sequences, a cloned HIV-1 genome fragment harbors a single sequence and is therefore ideal for validating the specificity of hybridization on microarray and for determining individual anti-tags that give rise to hybridization background. Four cloned fragments representing sequences of subtypes C (2 clones), CRF_02/AE, and A2 were used for validation. The agreement between SNP profiles from microarray versus Sanger sequencing was perfect for 2 plasmids containing subtype C1 and A2 sequences (0720235-C1 and 070510-A2). For another plasmid harbouring a subtype CRF_02/AE insert (6017225-AE2) 88% (22/25) agreement was reached and 72% (18/25) for a subtype C1 insert (072073-C1). As all plasmid inserts were derived from patients in Switzerland, they likely represented variants with substantial sequence deviation from our extension primer sequences, which had been optimized for Tanzanian subtypes. Overall the concordance between both typing methods was 90%, the data for each SNP tested is shown in **supplementary Table 2**.

Next, our HIV-1 SNP array was validated by genotyping 102 field samples from HIV-1 patients from Ifakara, Tanzania. The agreement between microarray and Sanger sequencing was 92.7% (2363/2550) (**supplementary Table 2**). Extension primers of 5 SNPs had a concordance of 100%. 14 SNPs had a concordance of >90% and only one

SNP had a concordance of <70% compared to the respective SNPs determined by Sanger sequencing. The SNPs with 100% agreement were: M41L1, L74V, V75I, T215FY and K219E. The SNPs with between 90% - 99% agreement were: M41L, D67E, K70R, T215FY1, T215FY2 and L74I (99%), K219Q (98%), L100I, Y181C and M184V (96.1%), M184I (95.1%), G190A and K103N (93.1%). The SNPs with <90% agreement were: D67N (87.3%), Y188L1 (86.3%), K219N (78.4%), K65R (76%), Y188I2 (71.6%) and L210W (62.7%). The L210W SNP principally performed poorly compared to all other SNPs. The sequence alignment of all 102 Tanzanian samples tested sequences revealed a number of mismatches with the designed extension primer in samples, which failed to produce a signal. Thus, the extension primer for SNP L210W was located in a region of considerable natural polymorphism (**supplementary Table 3**), and we failed to design a reliable primer, which would harbour less than three mismatches with any of the sequences. To address the polymorphism around SNP L210W, 8 related extension primers reflecting the possible sequences were designed for this SNP alone, but still only 62.7% of samples yielded a signal with this complex array. Primer mismatches are likely to be the main reason for missing data for false positive signals at this site. Another SNP, Y181C, initially also underperformed, but we re-designed a set of 5 representative extension primers for this SNP which yielded correct signals for 96% of samples.

DISCUSSION

The reliable microarray-based SNP typing approach, originally developed and in use for genotyping drug resistance markers of the malaria parasite *Plasmodium falciparum*

(Crameri *et al*., 2007) was now applied for genotyping DR-SNPs in the reverse transcriptase gene of HIV-1. The high mutation rate and genetic diversity were the toughest hurdles for a solely hybridization-based test. Our approach tried to overcome this challenge by (i) performing allelic discrimination by applying the mini-sequencing principle, and (ii) adapting the primers to the regionally prevailing viral strains. Our microarray hybridization was limited to perfectly base-paired tags and anti-tags, and we opted for adequate primer binding to diverse field isolates. Limitations in homology between sample and primers are the major challenges in developing a genotyping chip for HIV.

The overall aim of this development was to provide a cost-effective alternative to classical sequencing for resource poor setting, by typing only the minimal essential SNPs for the regionally available therapies.

On HIV-1 samples from Tanzania and with primers for 25 SNPs we were able to correctly identify 92% of all data points. As perfectly optimized sequences performed very reliably and with high sensitivity, we attributed suboptimal performance and missing data for certain mutations to mismatches of sample to primer. When we investigated the SNPs that had failed to produce signal we found that >3 mismatches within an extension primer was detrimental to hybridization and detection, while <3 primer mismatches located in the central position of the primer did not greatly affect PCR efficiency. As was known, also primer mismatches located within 4 nucleotides from the 3` terminus compromised PCR efficiency.

When disregarded (preliminary) extension primers, which in none of the experiments yielded a signal (i.e. 187/2550 SNPs), the concordance of microarray SNP profiles with Sanger sequencing reached 98%. To limit mismatching for distant HIV isolates we designed additional primers specific for some resistance positions. Since there was space for additional primers (due to our design for a limited number of reverse transcriptase mutations) we chose to use up to 5 primers for difficult positions or to incorporate wobbles at polymorphic positions: E.g. the design of five specific primers for Y181C yielded 96.1% success, and with four primers for K103N we obtained a signal in 93.1%. Only for SNP K210W despite the design of eight different extension primers we failed to reliably produce typing results for this SNP which resides within a hypervariable region of the gene.

A major task in the further development of this microarray will be to reduce the number of missing data, achievable by designing additional extension primers or by optimizing primer-annealing conditions, i.e. annealing temperature, duration, or salt concentration in the multiplex extension reaction. The advantage of a microarray with spotted anti-tags is that it is very flexible and allows addition of more or new tagged extension primers to the reaction for new mutations or new drug-resistance patterns.

This SNP-Microarray was developed primarily for population studies, e.g. for determining the prevalence of transmitted DR-SNPs, or for identifying reasons for treatment failure. For such questions, the deliberately limited set of informative SNPs genotyped by this method is adequate. For other research questions or for individual

diagnosis Sanger sequencing is more advantageous, as complete sequence information is gained.

CONCLUSION

Genotyping by microarray yielded a good agreement with Sanger sequencing in 102 field samples from Tanzania and resulted in correct base calling for 92% of SNPs. To generate complete DR haplotypes suitable for diagnostic purposes, further optimization will be required. Given the simplicity of its use, and a fairly short processing time to results, this hybridization-based microarray has demonstrated potential for monitoring resistance mutations in population-wide studies. The SNP-specific primers designed for this study may be equally useful for developing other hybridization-based genotyping tools, such as SNP typing on a bioplex platform or probe-based PCR applications.

ACKNOWLEDGMENTS

Dr. Johann Weber from the DNA array facility at the Center for Integrative Genomics, University of Lausanne is acknowledged for assisting with the printing of the slides.

This project was funded by the Swiss National Science Foundation grant no. IZ70Z0_131378/1.

Figure legends

Figure 1. HIV-1 SNP typing microarray. (A) Design and layout of microarray. Triplicates are depicted in different colours. T1, TC59, TC60, T61 and blnk denote Cy5-

spotting control-2, Cy5- spotting control-1, Cy3- spotting control, Cy3- pre-labeled hybridization control and printed buffer spots (blank), respectively. Some of the anti-tags were not in use in this experiment (T35, T36, T37, T48 and T52) and therefore utilized as negative control for calculating the cut-off. (B) Image of a HIV-1 SNP typing microarray after hybridization with a Tanzanian field sample. Each array consists of 14x14 spots. Cy5- spotting control (degenerated, indicated by green circles); Cy3-spotting control (red circles); Cy5-prelabeled hybridization control (blue circle); example of missing data (white circle).

Supplementary Figure 1. Validation of extension primers by PCR on cloned reverse transcriptase fragments and viral cDNA from Tanzania. Expected fragment sizes range from 56- 645 bp.

References

Bean P. New Drug Targets for HIV. Clinical Infectious Diseases (2005). 41(Supplement 1):S96–100.

Crameri A, Marfurt J, Mugittu K, Maire N, Regos A, Coppee JY, et al. (2007). Rapid Microarray-Based Method for Monitoring of All Currently Known Single-Nucleotide Polymorphisms Associated with Parasite Resistance to Antimalaria Drugs. J. Clin. Microbiol. 45(11):3685–91.

Fehr J, Glass T, Louvel S, Hamy F, Hirsch H, von Wyl V, et al. (2011). Replicative phenotyping adds value to genotypic resistance testing in heavily pre-treated HIV-infected individuals - the Swiss HIV Cohort Study. Journal of Translational Medicine.9(1):14.

Grant PM, Zolopa AR. The use of resistance testing in the management of HIV-1-infected patients. Current opinion in HIV and AIDS (2009).4(6):474–80.

Hirsch MS, Günthard HF, Schapiro JM, Vézinet FB, Clotet B, Hammer SM, et al. (2008). Antiretroviral Drug Resistance Testing in Adult HIV-1 Infection: 2008 Recommendations of an International AIDS Society-USA Panel. Clinical Infectious Diseases. 47(2):266–85.

Kaderali L, Deshpande A, Nolan JP, White PS (2003). Primer‐design for multiplexed genotyping. Nucleic Acids Research. 31(6):1796–802.

MacArthur R.D. (2009). Understanding HIV Phenotypic Resistance Testing: Usefulness in Managing Treatment-Experienced Patients. AIDS Rev. 11(4):223–30.

Mossdorf E, Stoeckle M, Mwaigomole E, Chiweka E, Kibatala P, Geubbels E, et al. (2011). Improved antiretroviral treatment outcome in a rural African setting is associated with cART initiation at higher CD4 cell counts and better general health condition. BMC Infectious Diseases.11(1):98.

Shafer RW, Kantor R, Gonzales MJ. (2000). The Genetic Basis of HIV-1 Resistance to Reverse Transcriptase and Protease Inhibitors. AIDS Rev. 2(4):211–28.

Syvänen A. (1999). From gels to chips: “Minisequencing” primer extension for analysis of point mutations and single nucleotide polymorphisms. Human Mutation. 13(1):1–10.

468 **Table 1**. Composition of Extension Primer and ddNTP reaction mixes

469

Extension Primer Mix 1	Extension Primer Mix 2
T-2-A-M41L ATTTTTGAAATTTTTCCTTCCTTTTCCA	T-1-A_K103N CCCACATCCAGTACTGTCACTGATTT
T-3-S_K65R TATAACACTCCAGTATTTGCCATAAAAA	T-5-A_D67E AAATCTACTAATTTTCTCCACTTAGTACT
T-4-A_D67N ATCTACTAATTTTCTCCACTTAGTACTGT	T-8-A_L74V TCTTTTATTGAGTTCTCTGAAATCTACTA
T-6-A_K70R TCCCTGAAATCTACTAATTTTCTCCACT	T-10-A_L100I AGTACTGTCACTGATTTTTTCTTTTTTA
T-7-S_K70R ATTTGCCATAAAAAAGAAGGACAGTACTA	T-13-S_Y188L2 ATCTATCAATACATGGATGACTTGT
T-9-A_V75I AGTTCTTTTATTGAGTTCTCTGAAATCTA	T-15-S_G190A TCTATCAATACATGGATGACTTGTATGTA
T-11-S_Y181C TAGAGCACAAAATCCAGAAATAGTTATCT	T-16-S_L210W AGAGGAGTTAAGAGCACATCTAT
T-12-A_Y188L1 TGCCCTATTTCTAAGTCAGATCCTAC	T-18-S_T215FY1 TAAGAGGACATCTATTGAGGTGGGGATTT
T-14-A-G190A CTCTATGCTGCCCTATTTCTAAGTCAGAT	T-19-A-K219Q ATGGAGGTTCTTTCTGATGTTTYT
T-17-A_T215FY1 TTCTGATGTTTCTTGTCTGGTGTG	T-20-S_M41L AGTGACAGTACTGGATGTGGGGG
T-23-A_T215FY2 TTCTTTCTGATGTTTCTTGTCTGGTGTG	T-21-S_T215FY AGAGGACATCTATTGAGGTGGGGATTTA
T-24-A_L74I CTTTTATTGAGTTCTCTGAAATCTACTA	T-22-A_T215FY TTCTGATGTTTTTTGTCTGGTGTG
T-28-S_184I AACCCAGAAATAGTTATCTATCAATATAT	T-27-A_K219N GAAATGGAGGTTCTTTCTGATGTTT
T-29-A_184V TAAATCAGATCCTACATACAAGTCATCCA	T-31-S_L74V AAGGACAGTACTAAGTGGAGAAAA
T-30-S_L74I AAGAAGGACAGTACTAAGTGGAGAAAA	T-33-S_L100I GGATACCACACCCAGCAGGG
T-32-S_L75I AAGGACAGTACTAAGTGGAGAAAATTA	T-33.2-S-L100I GGATACCACACCCAGCGGGG
160L74I1.1 ATTGAGTTCCCTGAAATCTACTA	T-34-S-K103N ACACCCAGCAGGGTTGAAAAAGAA
78M184V1.1 AGATCCTACATACAAATCATCCA	17M41L1.1 TTTGTAATTTTTCCTTCCTTTTCCA
144K219E1.1 TGGGGATTTACCACACCAGAC	02K103N1.1 ACATCCAGTACTGTCACTGATTT
T-02.1-A-Y181C ACATACAAGTCATCCATATATTGA	130L210W1.1 TCTGGTGTGGTAAATCCCCATTTI
T-38-A-Y181C ACATACAAGTCATCCATGTATTGA	T-39-S-L210W AGAGGAGTTAAGAGCACAYTTAT
T-40-A-Y181C ACATACAAGTCATCCACATATTGA	T-38-S_L210W AGAGGAGTTAAGAGCACATCTST
T-45.1-A-Y181C ACATACAAGTCATCCACATATTGA	T-41-A-L210W TCTCGTCTGGAGTGAAAAATCCCCATTTT
T-45-A-Y181C CATACAAGTCATCCACATATTGG	T-42-S-L210W AGAGGAGTTAAGAGCWCACCTAT
T-50-S-L100I GGATACCACACCCAGCAGGI	T-43-S-L210W AGAGGAGTTAAGAGCWCATCTAT
	T-44-A-L210W TCTTGTCTGGTGTGGTAAATCCCCATTTC
	T-46-S-K219Q TGGGGATTTACCACACCAGAI

T-47-A-L210W TTTTGTCTGGTGTGGTAAACCCCCACTTC

T-49-A-L210W TCTTGTCTGGTGTGGTAAATCCCCACCTT

T-51-S- K219Q TGGGGATTTACCACACCAGAC

Dideoxy-dNTP Mix 1	**Dideoxy-dNTP Mix 2**
ddATP-CY3	ddUTP-CY3
ddCTP-CY3	ddCTP-CY3
ddGTP-CY5	ddATP-CY5
ddUTP-CY5	ddGTP-CY5

470

471

472

473

474

475

476

477

478

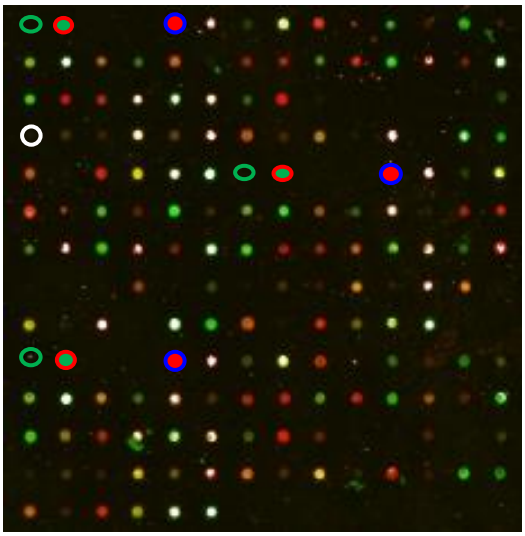


Fig. 2 Fig. 3

Figure 1. A B

485 **Supplementary Table 1**. List of extension primers, tags and anti-tags and spotting and

486 hybridization controls.

487

Extension Primer (5´→3´)	Tag (5´→3´)	Anti-Tag (5´→3´)
T-1-A_K103N CCCACATCCAGTACTGTCACTGATTT	T-1 GGTTCCCGATTTATCGATCCC	AT-1 GGGATCGATAAATCGGGAACC
T-2-A-M41L ATTTTTGAAATTTTTCCTTCCTTTTCCA	T-2 CATGTGGTACAATGGAACAGCTA CT	AT-2 AGTAGCTGTTCCATTGTACCA CATG
T-3-S_K65R TATAACACTCCAGTATTTGCCATAAAAA	T-3 TCAGGGAACTTCGATGCTGC	AT-3 GCAGCATCGAAGTTCCCTGA
T-4-A_D67N ATCTACTAATTTTCTCCACTTAGTACTG T	T-4 GACTGACCCGCTTGAGTTAGT	AT-4 ACTAACTCAAGCGGGTCAGT C
T-5-A_D67E AAATCTACTAATTTTCTCCACTTAGTAC T	T-5 GTTCAATCAGAAAACACCTGCGG	AT-5 CCGCAGGTGTTTTCTGATTG AAC
T-6-A_K70R TCCCTGAAATCTACTAATTTTCTCCACT	T-6 CTGCAAGCAGGTTGTGCTCT	AT-6 AGAGCACAACCTGCTTGCAG
T-7-S_K70R ATTTGCCATAAAAAAGAAGGACAGTAC TA	T-7 GGCGGTTCATGGAATTCCC	AT-7 GGGAATTCCATGAACCGCC
T-8-A_L74V TCTTTTATTGAGTTCTCTGAAATCTACT A	T-8 GTCCTACGTCGAGTAGAGAAAGT C	AT-8 GACTTTCTCTACTCGACGTAG GAC
T-9-A_V75I AGTTCTTTTATTGAGTTCTCTGAAATCT A	T-9 CATTTGCGTTTCTCTGGGTAATGC	AT-9 GCATTACCCAGAGAAACGCA AATG
T-10-A_L100I AGTACTGTCACTGATTTTTTCTTTTTTR	T-10 CCTGTCGGGAGCAGTACA	AT-10 TGTACTGCTCCCGACAGG
T-11-S_Y181C TAGAGCACAAAATCCAGAAATAGTTAT CT	T-11 ATCTACTACCACCTCCAACGG	AT-11 CCGTTGGAGGTGGTAGTAGA T
T-12-A_Y188L1 TGCCCTATTTCTAAGTCAGATCCTAC	T-12 GGGCGGACTACATCGAAATTACC	AT-12 GGTAATTTCGATGTAGTCCG CCC
T-13-S_Y188L2 ATCTATCAATACATGGATGACTTGT	T-13 CCGAAACAACGCAGAACTCAC	AT-13 GTGAGTTCTGCGTTGTTTCG G
T-14-A-G190A CTCTATGCTGCCCTATTTCTAAGTCAG AT	T-14 CTCTCCACAGTGCAGCGA	AT-14 TCGCTGCACTGTGGAGAG
T-15-S_G190A TCTATCAATACATGGATGACTTGTATGT A	T-15 TGGCCTTGTGAATCCACCC	AT-15 GGGTGGATTCACAAGGCCA
T-16-S_L210W AGAGGAGTTAAGAGCACATCTAT[1]	T-16 CGAAAAACCACGCCGTATTTCA	AT-16 TGAAATACGGCGTGGTTTTTC G
T-17-A_T215FY1 TTCTGATGTTTCTTGTCTGGTGTG	T-17 TCACTTACGACCGTTTTGTCTACA	AT-17 TGTAGACAAAACGGTCGTAA GTGA
T-18-S_T215FY1 TAAGAGGACATCTATTGAGGTGGGGAY TT	T-18 GAGAGGCATGCGTTTCACG	AT-18 CGTGAAACGCATGCCTCTC
T-19-A-K219Q ATGGAGGTTCTTTCTGATGTTTYT	T-19 GACCGGCAATTCGTTATCCAC	AT-19 GTGGATAACGAATTGCCGGT C
T-20-S_M41L AGTGACAGTACTGGATGTGGGGG	T-20 GTCAAATTCGACAGCTGGAAGG	AT-20 CCTTCCAGCTGTCGAATTTGA C
T-21-S_T215FY AGAGGACATCTATTGAGGTGGGGATTT A	T-21 GAAGCCGTCTCTGTTGTTTTCC	AT-21 GGAAAACAACAGAGACGGCT TC
T-22-A_T215FY	T-22 CAGAGATCCATTGGCGCGT	AT-22

TTCTGATGTTTTTTGTCTGGTGTI		ACGCGCCAATGGATCTCTG
T-23-A_T215FY2	T-23	AT-23 CTCGAAGTTGGGTCATTATG
TTCTTTCTGATGTTTCTTGTCTGGTGTY	CGCATAATGACCCAACTTCGAG	CG
T-24-A_L74I		AT-24
CTTTTATTGAGTTCTCTGAAATCTACTA	T-24 GCTGCCGGCTATTTTTGGAG	CTCCAAAAATAGCCGGCAGC
		AT-27
T-27-A_K219N	T-27	GTTAAACATAAACCTTCTCGG
GAAATGGAGGTTCTTTCTGATGTTT	CCCCCGAGAAGGTTTATGTTTAAC	GGG
T-28-S_184I		
AACCCAGAAATAGTTATCTATCAATATA		AT-28
T	T-28 AGCCTCGGGTCTACATCGT	ACGATGTAGACCCGAGGCT
T-29-A_184V		
TAAATCAGATCCTACATACAAGTCATC		AT-29
CA	T-29 CAGCAGTCCGATGCCTGG	CCAGGCATCGGACTGCTG
		AT-30
T-30-S_L74I	T-30	GGCTAGCTAAAAGGTCTAGG
AAGAAGGACAGTACTAAGTGGAGAAAA	CGCCTAGACCTTTTAGCTAGCC	CG
T-31-S_L74V		AT-31
AAGGACAGTACTAAGTGGAGAAAA	T-31 GGAGCTTTTGCTGTTCGGTC	GACCGAACAGCAAAAGCTCC
		AT-32
T-32-S_L75I	T-32	TGGTCAATAGTATGTCATACC
AAGGACAGTACTAAGTGGAGAAAATTA	CGGGGTATGACATACTATTGACCA	CCG
T-33-S_L100I		AT-33
GGATACCACACCCAGCAGGG	T-33 GTTGGCGGGTTATTACAGGG	CCCTGTAATAACCCGCCAAC
T-34-S-K103N		AT-34
ACACCCAGCAGGGTTGAAAAAGAA	T-34 TGCGATTGTATACCCGCTCC	GGAGCGGGTATACAATCGCA
		AT-35 GCGTAAATCATACGCCTGGG TC[2]
		AT-36 ACGCGTTACGTTAGAGATAA GGCTA[2]
		AT-37 GCCTCCACCCTTCTCAAGAA TA[2]
T-38-S_L210W		
AGAGGAGTTAAGAGCACATCTsT[1]		AT-38
	T-38 TTTCCGGATTCACCCGTACC	GGTACGGGTGAATCCGGAAA
T-39-S-Y181C		
TAGAGCACAAAATCCAGAAATAGTTAT		AT-39
wT	T-39 GATCGGACGACGCTTGGG	CCCAAGCGTCGTCCGATC
T-40-A-Y181C		AT-40
ACATACAAGTCATCCACATATTGA	T-40 TAGAGGAGGCGGGAGTTTTT	AAAAACTCCCGCCTCCTCTA
T-41-A-L210W		
TCTCGTCTGGAGTGAAAAATCCCCATT		AT-41
TT[1]	T-41	TGCACGAATTGTCATTCATTG
	AGCCAATGAATGACAATTCGTGCA	GCT
		AT-42
	T-42 GCACCACAGTCCGGTATTGC	GCAATACCGGACTGTGGTGC
		AT-43
T-43-A-K103N		GAAAAGTGGCCGTGTGTGAA
CCCACATCCAATACTGTTACTGACTT	T-43 TTTCACACACGGCCACTTTTC	A
T-44-A-L210W		
TCTTGTCTGGTGTGGTAAATCCCCATT		AT-44
TC[1]	T-44	CGTGACGCCACTAGTTCAAA
	TGTTTGAACTAGTGGCGTCACG	CA
T-45-A-Y181C		AT-45
CATACAAGTCATCCACATATTGG	T-45 GGTGATAGGCAACGAGGTCT	AGACCTCGTTGCCTATCACC
		AT-46
T-46-S-K219Q	T-46	AGCATCAAAGTCTAGGATCC
TGGGGATTTACCACACCAGAI	GGGGATCCTAGACTTTGATGCT	CC
T-47-A-L210W		
TTTTGTCTGGTGTGGTAAACCCCCACT		AT-47
TC[1]	T-47	GACAGAGACAGTCAGAGGAC
	CTTAGTCCTCTGACTGTCTCTGTC	TAAG

		AT-48 GACACACTTGTTGGACGCAA G[2]
T-49-A-L210W TCTTGTCTGGTGTGGTAAATCCCCACC TT[1]	T-49 GTGTTTGTCTACTTCGTGTGTGC	AT-49 GCACACACGAAGTAGACAAA CAC
T-50-S-L100I GGATACCACACCCAGCAGGI	T-50 ATGGAACCTATAATCTAGGATGGC G	AT-50 CGCCATCCTAGATTATAGGTT CCAT
T-51-A-L210W TTTTGTCTGGGGTAGTCAATCCCCAGC TC[1]	T-51 TCGTATAAGTCACGTTCTCCTTGG	AT-51 CCAAGGAGAACGTGACTTAT ACGA
		AT-52 CATTACTCCCTCCCGTCATGT [2]
17M41L1.1 TTTGTAATTTTTCCTTCCTTTTCCA	17.1 CAACATCATCACGCAGAGCATCAT T	17.1 AATGATGCTCTGCGTGATGA TGTTG
160L74I1.1 ATTGAGTTCCCTGAAATCTACTA	160.1 CCACGTACTGTCCGGAATACACG AC	160.1 GTCGTGTATTCCGGACAGTA CGTGG
02K103N1.1 ACATCCAGTACTGTCACTGATTT	02.1 TGCCCCGTTGCCCCGTTGCCCCG T	02.1 ACGGGGCAACGGGGCAACG GGGCA
75Y181C1.1 CATACAAGTCATCCATATATTGA	75.1 TAACACAAGAGCAGCTTGAGGAC G	75.1 CGTCCTCAAGCTGCTCTTGT GTTA
78M184V1.1 AGATCCTACATACAAATCATCCA	78.1 ACAGCCTCGCAGATGACGAATCA TT	78.1 AATGATTCGTCATCTGCGAG GCTGT
130L210W1.1 TCTGGTGTGGTAAATCCCCATTTI[1]	130.1.1 TACCAACTGTATGCGCATGTGCAC C	130.1.1 GGTGCACATGCGCATACAGT TGGTA
144K219E1.1 TGGGGATTTACCACACCAGAC	144.1.1 TTCAGTGTATGACGACCAGAGCG TT	144.1.1 AACGCTCTGGTCGTCATACA CTGAA
	[Cy3]AGAAGATGCCTAGTATATG	AT-61 CATATACTAGGCATCTTCT [Cy5]ATGCAACCATCAAGT- [AmC7~Q] [Cy3] GCTCAGCTGTATTAGAA- [AmC7~Q]

[1] Extension primers never giving a signal or with inconsistent performance

[2] Five additional anti-tags were designed and spotted to permit future use; these were utilized for quantification of background hybridization)

Supplementary Table 2. Comparison of microarray with direct sequencing using field samples and cloned DNA fragments.

SNP	M41L	K65R	D67N	D67E	K70R	L74V	V75I	L1001	Y181C	Y188L1	Y188L2	G190A	G190A	L210W	T215FY1	T215FY1	K219Q	T215FY	L741	K219N	M184I	M184V	K103N	M41L1	K219E	Agreemen
Field Samples																										
ET33	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	22
ET34	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	23
ET36	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	24
ET37	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	22
ET39	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	22
ET44	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	24
ET48	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	22
ET50	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	24
ET53	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	23
ET56	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
ET59	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	23
5510266	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
5510423	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	22
5510599	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	21
5510305	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	22
5510377	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	20
5510376	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
5510388	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	23
5511611	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
5410055	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
5510060	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	23
5510184	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	21
5510603	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	23
FTM	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	22
5510508	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	22
5510026	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	24
5510817	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25

5510236	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24
5510718	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	22
5510504	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	22
5510075	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
5510270	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	21
5510135	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	23
5514370	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	1	1	1	1	0	1	1	0	1	18
HLM5	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	23
HLM 6	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	21
HLM 7	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	1	1	21
HLM 13	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	23
HLM 17	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	23
HLM 23	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	21
HLM 25	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM 27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM29	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	22
ET 2	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	0	1	1	1	1	20
HLM 32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 35	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM 36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 69	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 71	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24
HLM 73	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 74	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM76	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	23
HLM 77	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24
HLM 78	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24
HLM 79	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 80	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 83	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	23
HLM 86	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM 87	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24
HLM 88	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	24

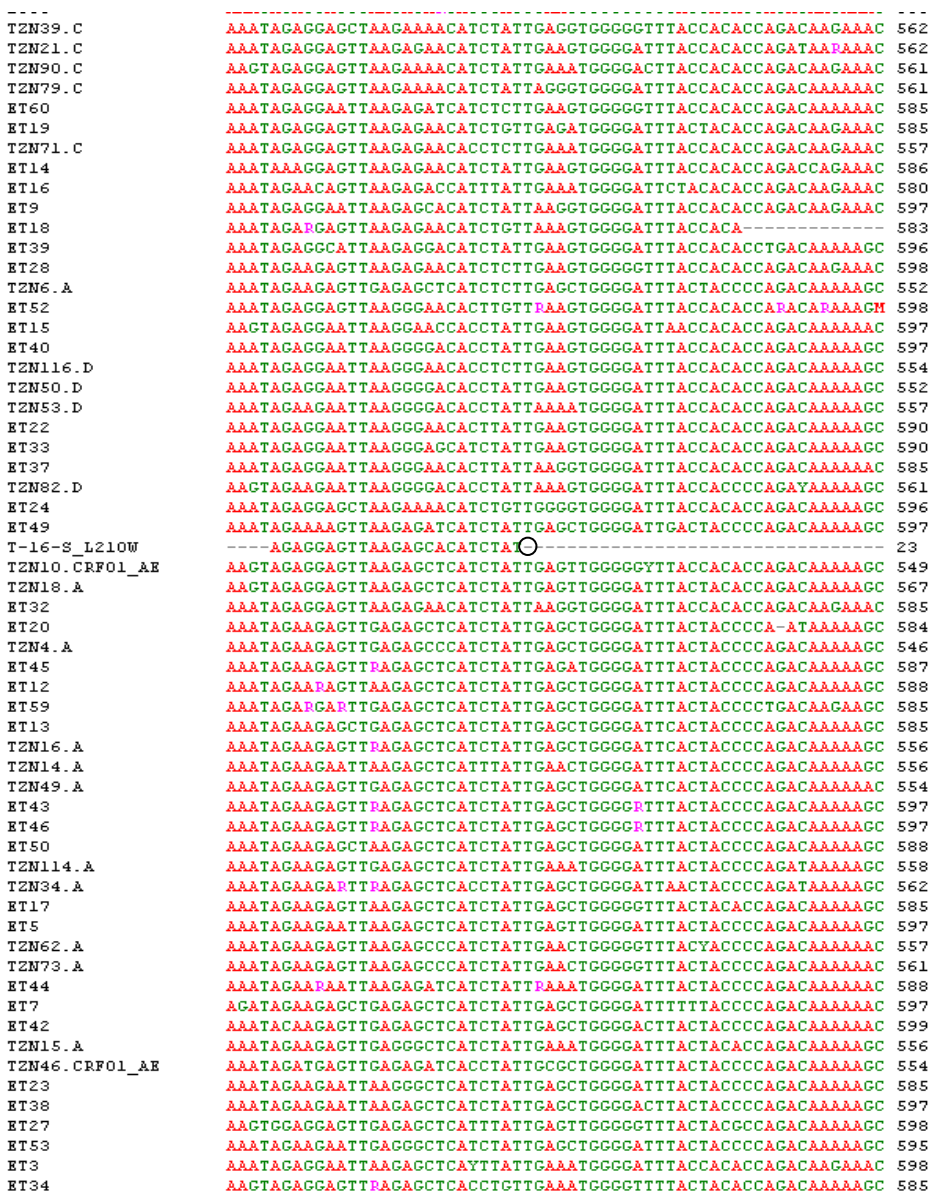
ET 7	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	0	1	1	0	1	1	19
ET 7	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1	19
5514352	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	20
5510255	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	22
5511110	1	1	0	1	1	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	1	18
5514391	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23
HLM-89	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-91	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-92	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-93	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	23
HLM-172	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-173	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-175	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-176	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	24
HLM-177	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-178	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-180	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	23
HLM-181	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-182	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-185	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-258	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
HLM-259	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	22
HLM-260	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	22
HLM-261	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	22
HLM-262	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	23
HLM-263	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-264	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	23
HLM-265	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	24
HLM-267	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-182	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-271	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23
HLM-267	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-182	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
HLM-271	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23
05510783-	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24

A																									
5510075	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	23
ET 7	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	0	1	1	0	1	1	19
HLM 88	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	24
HLM 80	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
07510783-B	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	22
% agreement	99	76	87	99	99	100	100	96	96	86	72	93	93	63	99	99	98	100	99	78	95	96	93	100	100	92.7
agreement	101	77	89	101	101	102	102	98	98	88	73	95	95	64	101	101	100	102	101	80	97	98	95	102	102	2363
Cloned RT Fragments																										
0720235-C1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
6017225-AE2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	22
072073-C1	1	0	1	1	1	1	1	0	0	0	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	18
070510-A2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
agreement	4	3	4	4	4	4	4	3	3	3	3	3	4	3	4	4	4	4	3	3	3	4	4	4	4	90

497

498

Supplementary Table 3. Location of L210W SNP in the Alignment of Ifakara RT sequences



Circle represent the location of the L210 SNP, which is preceded by a region of high polymorphism